

OPERATING SYSTEMS INC WOODLAND HILLS CA
ASSOCIATIVE FILE PROCESSOR.(U)

ASSOCIATIVE FILE PROCESSOR: (U)
FEB 80 R L PICKLER, D K WHELCHER
OSI-R79-031

DTIC USERS

UNCLASSIFIED

RADC-TR-80-23

1 of 1
A0
AQ83487

END
DATE
FILMED
5 80
DTIC

5 80
DUE

RADC-TR-80-23 ✓
Final Technical Report
February 1980

12
LEVEL II



ASSOCIATIVE FILE PROCESSOR

Operating Systems, Inc.

R. L. Pickler
D. K. Wheelchel

DTIC
ELECTED
APR 25 1980
S D C

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

ROME AIR DEVELOPMENT CENTER
Air Force Systems Command
Griffiss Air Force Base, New York 13441

DOC FILE COPY

ADA 083487

80 4 25 007

This report has been reviewed by the RADC Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

RADC-TR-80-23 has been reviewed and is approved for publication.

APPROVED:



PATRICIA M. LANGENDORF
Project Engineer

APPROVED:



HOWARD DAVIS
Technical Director
Intelligence & Reconnaissance Division

FOR THE COMMANDER:



JOHN P. HUSS
Acting Chief, Plans Office

If your address has changed or if you wish to be removed from the RADC mailing list, or if the addressee is no longer employed by your organization, please notify RADC (IRDA), Griffiss AFB NY 13441. This will assist us in maintaining a current mailing list.

Do not return this copy. Retain or destroy.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

1. REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM	
1. REPORT NUMBER RADC/IR-80-23	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) ASSOCIATIVE FILE PROCESSOR • /	5. TYPE OF REPORT & PERIOD COVERED Final Technical Report.		
6. AUTHOR(s) R. L. Pickler D. K. Whelchel	7. PERFORMING ORG. REPORT NUMBER OSI-851- R79-031	8. CONTRACT OR GRANT NUMBER(s) F30602-78-C-0133	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Operating Systems, Inc. 21031 Ventura Blvd Woodland Hills CA 91343	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62702F 45941603	11. REPORT DATE February 1980	
12. CONTROLLING OFFICE NAME AND ADDRESS Rome Air Development Center (IRDA) Griffiss AFB NY 13441	13. NUMBER OF PAGES	14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Same	
15. SECURITY CLASS. (of this report) UNCLASSIFIED		16. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A	
17. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.			
18. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) Same			
19. SUPPLEMENTARY NOTES RADC Project Engineer: Patricia Langendorf (IRDA)			
20. KEY WORDS (Continue on reverse side if necessary and identify by block number) ASSOCIATIVE FILE PROCESSOR ASSOCIATIVE FILE PROCESSOR software enhancement query generation Document retrieval			
21. ABSTRACT (Continue on reverse side if necessary and identify by block number) This report describes the work accomplished under Contract No. F30602-78-C-0133 for Rome Air Development Center (RADC). The objective of this contract was to support advanced software development for the Associative File Processor (AFP), developed by Operating Systems, Inc. (OSI). The AFP, designed and built by OSI, is a special purpose system utilizing the DEC PDP-11 family of computers. including (Cont'd)			

DD FORM 1473 EDITION OF 1 NOV 58 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

391974

117

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

Item 20 (Cont'd)

the AN/GYQ-21(V) processors and disk storage devices. It is designed to search large data bases (1-10 billion characters of on-line files) of unstructured free-text data for multiple, random queries within a reasonable time and at an affordable cost. The AFP and its current prototype software have been especially designed for intelligence applications.

The technical effort for this contract was directed toward general improvement of the AFP operation in several areas: increased user interface flexibility, additional utility programs, expanded functional capabilities, etc.

For the most part, the goals delineated by the contract have been met. This was accomplished by modifying existing AFP software to expand AFP capabilities and by developing new software modules for additional functions.)

Section 2 of this report provides a general description and background of the AFP. The enhancements provided under this contract are discussed in Section 3.

While the capabilities of the AFP have been significantly increased as a result of the work done under this contract, several other areas for improvement or increased flexibility of the AFP have been suggested. These areas are mentioned in Section 4.

Accession	✓
NTIS	
DDC TAB	
Unannounced	
Justified	
By	
Date	
Approved	
Dist	
A	

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

CONTENTS

SUMMARY	1
1. INTRODUCTION	3
2. DESCRIPTION OF ASSOCIATIVE FILE PROCESSOR (AFP)	5
2.1 General	5
2.2 AFP Hardware Configuration	7
2.3 AFP Software	7
2.4 Existing Functional Capabilities	15
2.5 General Operational Procedures	17
3. FUNCTIONAL ENHANCEMENTS UNDER CURRENT CONTRACT	21
3.1 General	21
3.2 Program to Collect GENSER Message Traffic	21
3.3 GENSER Message Collection	22
3.4 Document Editor Program Modifications	22
3.5 Synonym Dictionary Capability in Queries	24
3.6 Synonym Dictionary Listing Program	26
3.7 Multi-User Capability on UNIVAC 1652	26
3.8 Distinguishing Among Responses to Multiple Queries	27
3.9 Allowing for Larger Number of Query Responses (Overflow)	28
3.10 Printing a Document on the Line Printer	28
3.11 Creation of an Editable Document	29
3.12 Concordance (Synonym Candidate) Listing Program	29
3.13 Deletion of Documents from Search File	30
3.14 Address List for Performing Subset Queries	31
3.15 Don't Care	32
3.16 Highlighting	33
4. AREAS FOR POSSIBLE FUTURE ENHANCEMENTS	34
4.1 General	34
4.2 Receipt of Real-Time Messages	34
4.3 Automated Dissemination	34
4.4 Incorporation of PDP-11/04 into AFP System	35
4.5 Improved Editing and Report Generation	35
4.6 Conversion to DEC IAS Operating System	37
4.7 Conversion to DEC UNIX Operating System	37
4.8 User Interface Enhancement	39
4.9 Improved Response Time	40
4.10 Search of a Mixed Format Data Base	40
4.11 Use of AFP with ARPA Network	41
REFERENCES	42

EVALUATION

The Associative File Processor (AFP) is a potential device to support searching very large disk-resident computer files. This contract developed the software required to make the AFP fit into intelligence data handling application. This is important because many intelligence data handling applications are search-time limited.

Patricia M. Langendorf
PATRICIA M. LANGENDORF
Project Engineer

SUMMARY

This report describes the work accomplished under Contract No. F30602-78-C-0133 for Rome Air Development Center (RADC). The objective of this contract was to support advanced software development for the Associative File Processor (AFP), developed by Operating Systems, Inc. (OSI).

The AFP, designed and built by OSI, is a special purpose system utilizing the DEC PDP-11 family of computers, including the AN/GYQ-21(V) processors and disk storage devices. It is designed to search large data bases (1-10 billion characters of on-line files) of unstructured free-text data for multiple, random queries within a reasonable time and at an affordable cost. The AFP and its current prototype software have been especially designed for intelligence applications.

The technical effort for this contract was directed toward general improvement of the AFP operation in several areas: increased user interface flexibility, additional utility programs, expanded functional capabilities, etc.

For the most part, the goals delineated by the contract have been met. This was accomplished by modifying existing AFP software to expand AFP capabilities and by developing new software modules for additional functions.

Section 2 of this report provides a general description and background of the AFP. The enhancements provided under this contract are discussed in Section 3.

While the capabilities of the AFP have been significantly increased as a result of the work done under this contract, several other areas for

improvement or increased flexibility of the AFP have been suggested. These areas are mentioned in Section 4.

1. INTRODUCTION

The Associative File Processor (AFP) is a system designed and built by Operating Systems, Inc. which provides for high-speed searching of large, unstructured data bases. The AFP system comprises a hardware associative/parallel matching device (AXP), a CPU, a mass storage device with controller, and system software. Because of its ability to rapidly retrieve documents containing selected textual information from large data bases, the AFP has obvious implications for use in intelligence applications.

Contract No. F30602-78-C-0133, under the aegis of Rome Air Development Center (RADC), was issued to Operating Systems, Inc. to exploit the potential of the AFP for intelligence applications. In particular, the objective of this contract was to provide for advanced development of software for the AFP to enhance its usefulness as an intelligence tool.

This report describes the results of the effort expended under the referenced contract and the nature of the resulting enhancements to the AFP's operation.

Section 2 of the report describes some particulars of the AFP system and its general status prior to the work done under the referenced contract.

The tasks accomplished and the enhancements made to the AFP system software are discussed in Section 3.

Section 4 briefly discusses areas for possible future enhancements to the

AFP to render it more valuable as an aid in intelligence applications.

2. DESCRIPTION OF ASSOCIATIVE FILE PROCESSOR (AFP)

2.1 General

This section describes the state of development of the AFP prior to undertaking the effort outlined in the referenced contract.

The Associative File Processor is a fast, associative, disk file search system. The purpose of the AFP is to search a basically unstructured data base against multiple queries simultaneously, and to retrieve information where search criteria are met. What gives this search system its processing power is the Associative Crosspoint Processor (AXP), a hardware approach to file processing, effectively having the power of 1200 CPUs operating simultaneously. This parallel search unit performs searches independently of, but in concert with, the CPU. The AFP consists of the AFP software, an RSX-11D operating system, the AXP, a search disk and controller, and a PDP-11 series host computer with attendant peripherals.

The AFP software consists of the following three components:

- Search File Generation
- Query Generation and Document Retrieval
- Search Control

The Search File Generation software converts collected documents or messages into an AXP searchable disk file.

The Query Generation and Document Retrieval software allows a user to enter

queries for search criteria and then retrieve the documents which are satisfied by the queries.

The Search Control software compiles queries, loads and starts the AXP, and evaluates the queries for match and later document retrieval.

The AFP operates under the standard DEC RSX-11D operating system software. The RSX functions utilized by the AFP system are:

- Console Monitor (MCR)
- Text Editor (EDI)
- Peripheral Interchange Programs (PIP)
- Macro Assembler (MAC)
- Task Build (TKB)

The basic unit of searchable information for the AFP is a document, message, or any definable quantum of information. A search file is a collection of documents of textual information in a data base. To identify each document or basic information unit, boundary markers are placed at the beginning and end of the document. With these markers established, the AFP is able to simultaneously test user queries at lower levels of the basic information unit, such as words or phrases within the individual documents. The markers are sensed by the AXP during searching and the disk address corresponding to the beginning of document is then sent to the host CPU. When a search criterion has been satisfied (i.e., a query match found) the complete document can be retrieved.

2.2 AFP Hardware Configuration

An AFP system is constituted by combining a PDP-11 series computer with the appropriate peripherals. With the addition of an AXP, a Busrouter, and a searchable disk and controller to the basic computer configuration, an AFP system is configured. The AFP minimum hardware configuration (Figure 2-1) consists of:

- PDP-11 series computer with memory management
- 88K words of core or solid-state memory
- System disk, w/8K sectors
- 1 to 4 CRT's TTY compatible
- AXP
- Search disk and controller
- Magnetic tape unit
- Busrouter

2.3 AFP Software.

This section describes the Associative File Processor (AFP) software capabilities as they existed prior to this contract effort.

The Associative File Processor consists of software and utilities which provide system, user and data base support, while also providing considerable system flexibility for most applications.

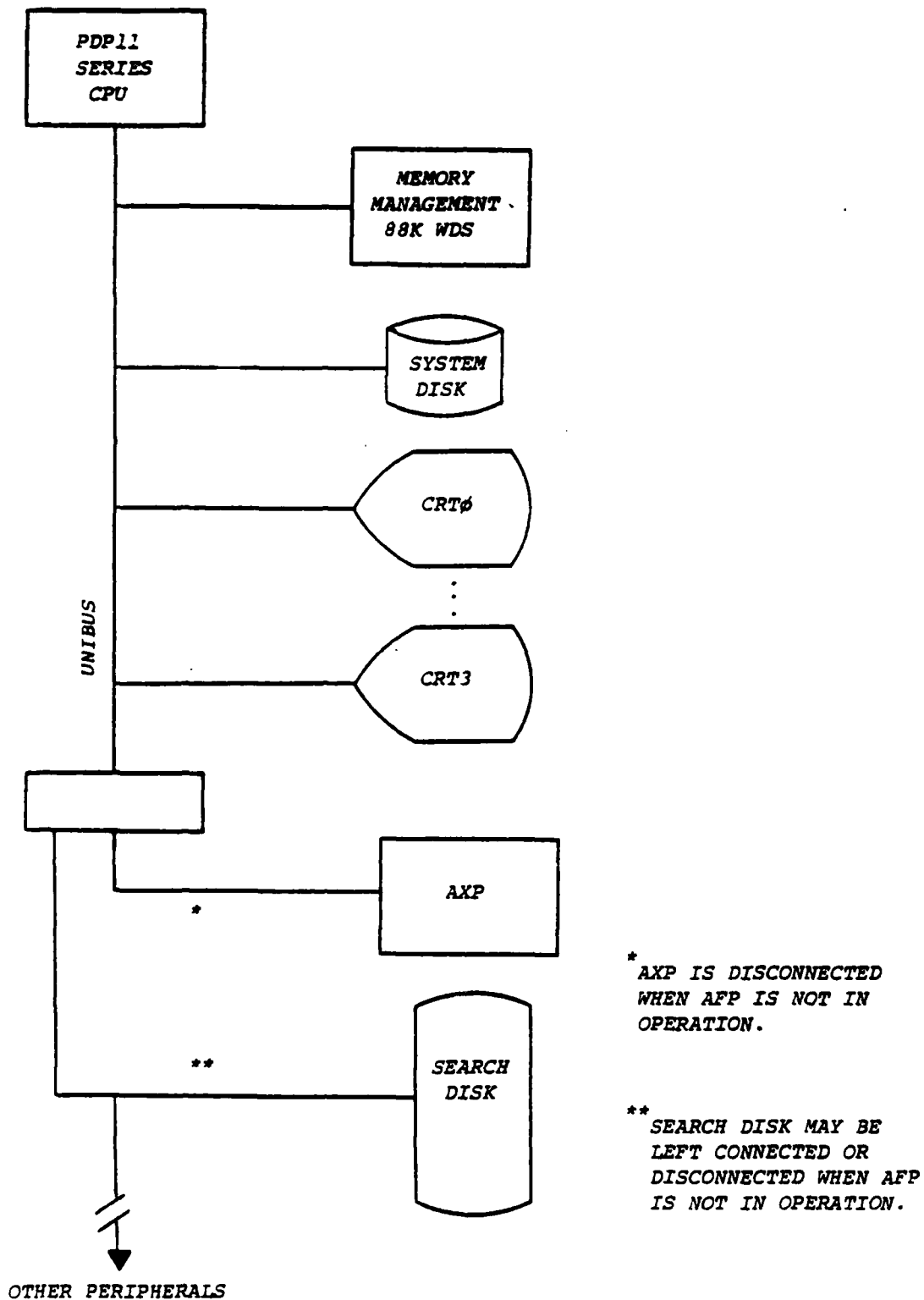


Figure 2-1. AFP Hardware Configuration

The capabilities provided by this software include user interfaces and applications programs for:

- the on-line creation of queries in a logical, English-like language,
- initiating high speed, full text searches of a large data base against multiple queries for several users,
- on-line message retrieval and review, and
- data base creation and update.

The AFP software is divided into the following functional groups:

- (a) Data Base Generation
- (b) Query Generation
- (c) Query Compilation
- (d) AXP Search Control
- (e) Query Resolution
- (f) Document Retrieval
- (g) System Diagnostics

The AFP software runs under the RSX-11D operating system, version 6.2 (which is upward compatible with the IAS operating system) and makes use of some of the system utilities for file creation, editing and manipulation. These include the following:

(a) Peripheral Interchange Program (PIP) for file transfer, renaming and deletion.

(b) Direct modification of file locations (and task images) (ZAP).

(c) File Dump Utility (DMP) for examining the ASCII, byte or word content of file blocks.

(d) File Comparison Utility (CMP) for comparing the contents of two files. The differences are listed.

(e) File line editor (EDI) for creating and editing RSX-11 files.

2.3.1 Data Base Generation Software. This software consists of search file disk formatting and document editing programs.

The disk formatting software (SEG) divides the search-disk file space into a number of pre-allocated, empty files of a pre-determined length and consisting of contiguous disk blocks. These files are called segments. The program that loads searchable files onto the search-disk is called the Document Editor (DOCEDI). The Document Editor is an interactive program which prompts the user for the various parameters required to create a data base. Different versions of this program are needed for the various data bases that are written to the disk; for example EDFBIS is required to read messages which are in the Foreign Broadcast Information Service format from magnetic tape to the disk.

The operational procedures for formatting a search-disk and for using the various versions of the Document Editor are described in the publication

entitled "FILE MANAGEMENT FOR AFP SYSTEMS, Concepts and User Instructions", Part Number UM131009. When a new data base is created on the search disk, the first available unused segment is accessed by the Document Editor Program, renamed to the desired file name, and subsequently overwritten with the new data. This process is repeated as often as necessary until the requisite number of data blocks are transferred to the search-disk.

During the process of transcribing documents from one medium to another (from tape to disk or disk to disk) some processing is performed on the data. For example, certain non-printing control characters, which might be interpreted as commands by the AFP, are stripped out.

2.3.2 Query Generation. User queries can be created on-line or from an indirect query file. The interface for the generation of queries is the Query Language Translator (QLT). The Query Language Translator is a multi-user program which interprets the user-entered query(s) and creates tables for use in query resolution.

Queries can consist of a combination of natural language and Boolean terms and phrases. The Boolean terms are enclosed in single quotes and are separated by logical AND's, OR's, NOT's and proximity indicators. A user's guide for the creation of queries is found in the publication "USER'S MANUAL FOR THE ASSOCIATIVE FILE PROCESSOR (AFP)."

2.3.3 Query Compilation. The query compilation function is controlled by the Search Monitor (SCH). The modules involved in this function are:

(a) COMPI

(b) PMAPBL

(c) COMP2

(d) MAC (RSX-11 Assembler)

(e) TKB (RSX-11 Task Builder)

The COMP1 program tabulates the queries resulting from the Query Generation process and creates a Query Expression Table for use by the PMAPBL task. This table serves as an outline for the query resolution rules.

The PMAPBL program organizes the query terms for loading into the AXP keyword memory. Pointer Index and Pointer Memory parameters are calculated, resulting in two outputs from this module: search parameters, which are to be loaded into the keyword memory, and memory mapping vectors for each term. The memory mapping vectors are used as inputs to the COMP2 program.

The COMP2 module merges the memory mapping vectors with the Query Expression Table and replaces each query term with a vector, resulting in a table which can be assembled and task-built.

The output of COMP2 is assembled and task-built to produce the final machine-processable Query Resolution Table.

2.3.4 Query Resolution. The query resolution process is controlled by the Query Resolution Task (QRT). Query expressions are compared to word matches encountered during the AXP search process. The word match statuses returned by the AXP are accompanied by memory vectors which identify the query terms found in the search. When QRT determines that a match has occurred between

a user's query and the terms in a document, pointers containing the document's disk address are sent to the document retrieval program.

2.3.5 Document Retrieval. The document retrieval process is controlled by the Document Retrieval Monitor (AMON). AMON is a multi-user task which allows up to four users¹ to retrieve and review documents from a CRT terminal. Documents may be reviewed serially, as they are queued for review. The user may also skip forward and backward in the queue. Pages within a document may be skipped forward and backward, as well.

Documents found by the Query Resolution Task to match a particular query are passed² to the Document Retrieval Monitor, where they are placed in a queue for the terminal from which the original query was created. Each queue entry, or node includes the address of the document on the search disk and the query identification of the query(s) satisfied by the document. The Monitor incorporates node management software whereby nodes may be drawn from a common pool and linked to other nodes corresponding to a particular user's retrieval queue.

The AMON Task also maintains separate I/O and status areas for each user, permitting asynchronous I/O.

The Document Retrieval Monitor employs two other routines for formatting and

1. Each user may enter up to twenty-five queries. Further query specifications are given in the Query Specification Table contained in the users' manual referenced previously.
2. Note that the document text is not passed to the Retrieval Monitor, but the beginning and ending address of the document on the search disk, the user terminal(s) for which the document is to be queued and the alphanumeric identification of the query(s) satisfied by the document.

displaying documents on the CRT screen, these are:

(a) CRT Library Routine (CRTSUB) and

(b) Format Module (FRM).

CRTSUB is a library routine which accepts text lines from AMON and builds a complete CRT page for display. FRM composes documents to fit the dimensions of the user's CRT screen.

2.3.6 System Diagnostics. The system diagnostics consist of a set of confidence level tests and a diagnostic exerciser. The tests and exerciser are discussed in the publication, "ASSOCIATIVE FILE PROCESSOR (AFP) MAINTENANCE MANUAL", Part number MM131005 V00R00.

The confidence level tests are a set of standard queries run against a base-line data base. The search hardware and query resolution software are evaluated by matching the number of hits attained for each of these confidence tests to the expected results. The retrieval software is verified by retrieving documents in a predetermined sequence and comparing the identification number of the particular documents retrieved against the expected documents.

If any one of these tests is not passed the Diagnostic Exerciser (INBST) is run. The exerciser isolates AXP hardware memory and I/O problems by systematically loading and reading hardware memory locations, and comparing the results to the inputs.

2.4 Existing Functional Capabilities

2.4.1 Hardware Capabilities. The major components of the Associative File Processor, AFP, are: a host PDP-11/34 or 45 or 70 series host processor, with 88K words of memory, or more, a system disk and controller, a large capacity disk and controller for archival storage and retrieval, an Associative Crosspoint Processor, AXP, a Busrouter, user terminals, appropriate interfaces, an RSX-11D Operating System and software utilities, and AFP Support Software.

The Associative File Processor supports up to four user terminals,³ which may be CRT's, or other teletype-compatible terminals. Query creation and text search are initiated by MCR commands entered at the keyboard and document retrieval is controlled by entering alphanumeric mnemonics.

2.4.2 Functional Capabilities. Users may create queries, search data bases and retrieve documents independently of one another. The AFP will simultaneously search multiple queries entered by one or more users; however, only one data base may be searched at a time.⁴

Queries may be created in either an on-line mode, using the Query Language Translator, which prompts the user for entries, or off-line using the RSX-11D editor. Queries may consist of a combination of Boolean expressions and natural language text. The Boolean expressions allowed are: AND, OR, NOT

3. This is a limitation placed on the number of terminals recognized by the AFP and is not a limitation on the number of terminals that the operating system may be sysgened to recognize.

4. A multi-user search is initiated when several users enter the search command within the same time frame--approximately ten seconds. The procedure is explained in the next subsection.

and WITHIN n, the latter expression being a proximity operator which allows windowing.

A single query can consist of up to 26 terms or contiguous-word phrases and up to 26 words per phrase. Single words of up to 15 characters are allowed; words of greater length are truncated. A maximum of 25 queries per user is allowed. The AXP will accept a maximum of 8,192 characters.⁵

A maximum of one hundred fifty documents may be retrieved by a single user. The search and retrieval processes are asynchronous. Document retrieval is initiated automatically during a document search. As soon as the first query match is made the document is placed in the appropriate user review queue and is displayed on that user's terminal. Subsequent documents may be reviewed by the user in any order. The user may page forward and backward within a document.

When the search process is completed a new search may be initiated against new queries by some users, while others may continue to review documents retrieved from the previous search, or create new queries.

The capability is provided for formatting search disk packs and for creating an archival data base on such packs from a message or document file stored on magnetic tape. Documents may also be added to an existing data file. The number of data bases which may reside on a given disk pack is dependent on the capacity of the disk pack and the size of the files to be stored on the pack.

5. The sum of all of the characters in every keyword in all queries for all users may not exceed this number for a single search.

2.5 General Operational Procedures

This subsection presents, as background information, a brief summary of the procedures for using the Associative File Processor. The description includes system startup procedure, creation of a query, initiation of a search and document retrieval. The detailed procedures for system preparation and startup are presented in the "ASSOCIATIVE FILE PROCESSOR PROGRAMMER'S MANUAL", Part Number AA-PROG-000-000-1. System operation, for query generation, searching and document retrieval, is detailed in the "USER'S MANUAL FOR THE ASSOCIATIVE FILE PROCESSOR (AFP)", Part Number AA-USERS-000-000-1. Search-disk formatting and search file generation and update are described in the "DOCUMENT EDITOR PROGRAMMER'S MANUAL", Part Number PM131007 VO1R00.

2.5.1 System Preparation and Startup. The RSX-11D System Generation command file must be revised to reflect the devices used by the AFP and the system Address Paging Register (APR) usage. Additional devices required for the AFP include the AXP hardware and the search disk. The system must be generated and the required standard system software must be installed. The general procedure outlined in the RSX-11 "SYSTEM GENERATION" manual should be followed.

The AFP software is then transferred from magnetic tape to the system disk and installation-dependent programs are modified. The AFP software is then assembled and task-built using command files which automatically initiate these processes for each task and move the various modules to pre-assigned user areas on the system disk. The AFP programs are installed and the sys-

tem is saved⁶. After the successful completion of system checkout the system is preserved on magnetic tape, providing a ready backup system. The system is normally saved with the disk and AXP handlers loaded and the search and system disks mounted, permitting the AFP to be used immediately after the system is booted.

2.5.2 Query Generation. Queries are generated using the Query Language Translator (QLT) task and may be created on-line or from an existing indirect file created with the RSX-11D editor.

The Query Language Translator is invoked by a user command entered at the terminal. QLT will respond with a prompt requesting that the user enter either: (a) an indirect file name, which contains a pre-composed query, or (b) a carriage return. A carriage return indicates that an on-line query is to be generated.

If an indirect file name is entered the Translator will access the file, interpret the query and produce a query verification listing on the terminal, giving a logical breakdown of the query as it was interpreted by the AFP. The user is then prompted to either accept or reject the query. If the user elects to accept the query (by responding yes followed by a carriage return), then he may enter additional queries, or terminate the session. If the query is rejected (by entering no followed by a carriage return), then he may re-enter the query.

6. Installed tasks are recognized by the operating system as run-ready tasks which may be called by the user, or automatically by other programs. Saving the system with the desired tasks installed will cause those programs to be permanently installed, obviating the need for re-installing them each time the system is booted.

An on-line query is processed the same as an indirect-file query, except that the user enters the query text at the terminal under QLT control and must specify the end-of-query to that task.

The query session is ended by a user command. The user may then proceed with a search, or perform other tasks; however, when a search is eventually initiated from that terminal, the query compilation software will accept the query file⁷ created during the most recent query session.

2.5.3 Search Initiation. An AFP search is initiated from the user's TTY or CRT terminal by typing-in the characters SCH followed by a carriage return. If multiple users wish to search the same data base simultaneously, they must input their search requests within ten seconds of the first user to initiate the search. The search software has a built-in ten second delay. After the delay is ended, the software polls its search request queue to determine which terminals have requested a search, then processes⁸ the queries for those users.

When the queries have been processed for query resolution and loaded into the AXP hardware memory, the user will be prompted for the name of the file to be searched. In the case of a multi-user search, the first terminal to initiate the search is the only terminal to be prompted for the search file name. The search file name is entered at the terminal and is followed by a

7. This refers to the query file created by Query Language Translator task from the user-entered query.

8. Processing, at this point, consists of assembling and building query tables. One table is loaded into the AXP memory and contains all of the key query words. Another contains the query logic and is used by query resolving software to resolve the key words in a document against the query logic.

carriage return, which starts the AXP search of the data base.

Documents satisfying the various user queries are queued-up for review at the originating terminals. The first document retrieved is automatically displayed at the user's terminal, while the search continues.

2.5.4 Document Retrieval. Document retrieval begins as soon as the first document is found to match a user's query. The first display retrieval is initiated automatically by the retrieval software. Subsequent display retrievals are at the user's command. Multiple user's may retrieve documents independently of one another.

At some point some users will want to create new queries, while others may wish to continue the retrieval and review process and still others may wish to search a new data base. These processes are mutually independent and may occur simultaneously, from different terminals.

Documents displayed at the terminal are preceded with a header containing information about the search status (SEARCHING or DONE), the number of documents that matched the user's query(s), how many are queued for review, the query identification(s) of the query(s) satisfied in the document, the position of the document in the queue (e.g., 1, 2, nth), number of pages in the document (terminal pages), and the number of the current page being displayed.

The user may page through the document, skip pages forward and backward, request the next document to be displayed, skip documents forward and backward, and exit the retrieval session.

3. FUNCTIONAL ENHANCEMENTS UNDER CURRENT CONTRACT

3.1 General

This section describes the enhancements added to the basic AFP software as a result of the effort expended under the referenced contract. The nature of each task defined by the contract is briefly described, as well as the task's outcome.

3.2 Program to Collect GENSER Message Traffic

The original search data base used by Operating Systems, Inc. for developing and testing the AFP consisted entirely of Foreign Broadcast Information Service (FBIS) messages. To more adequately demonstrate the usefulness of the AFP in an intelligence environment, however, a data base consisting of intelligence messages is more relevant. Consequently, this task was designed to provide a data base of GENSER messages for searching by the AFP.

Operating Systems, Inc. had already developed a program for collecting GENSER messages for the NMIC MSS subsystem prior to this effort. For this task, however, the program had to be modified and expanded to collect messages from the NMIC 5-day file and write them to magnetic tape in a format suitable for processing by the AFP search file edit program (DOCEDI).

This task was accomplished, and the GENSER messages were collected. (See Section 3.3).

3.3 GENSER Message Collection

For this task, Operating Systems, Inc. personnel in our Arlington, Virginia, office collected GENSER messages from the 5-day file at the NMIC PAS computer facility in Washington, D. C. These messages were collected using the GENSER collection program discussed in Section 3.2.

Although it was our original intention to collect about 20,000 unclassified messages for the GENSER data base, it was later determined that the capacity of the NMIC 5-day file is insufficient to hold that many unclassified messages. Consequently, we collected all the unclassified messages that were available on the 5-day file and, using a special feature of the Document Editor program, replicated the collected messages to produce a data base sufficiently large to meet the needs of the contract.

3.4 Document Editor Program Modifications

The purpose of the Document Editor program (DOCEDI) is to accept data from either disk or magnetic tape and process the data to create a file which is suitable for searching the AFP.

For the current project, modifications were made to DOCEDI to provide significant improvements and extend its capabilities in the following areas:

- Creating RSX-compatible search files
- Accepting GENSER data
- Appending documents to search files

These areas are described below.

3.4.1 Creating RSX-Compatible Search Files. Prior to this project, the AFP search files created by DOCEDI were not compatible with the file management system of the PDP11 RSX operating system. This incompatibility conducted to several problems. For one thing, AFP search files are normally larger than 32K disk sectors. But since the RSX file management system restricts file length to 32K sectors, the AFP search files could not be recognized by the RSX file management system. This meant that the normal file services such as adding, deleting, editing, and printing files were not available for the AFP search files.

For this task, changes were made to DOCEDI such that the search files it creates for the AFP are now RSX-compatible and can utilize the normal file services of the RSX file management system.

3.4.2 Accepting GENSER Data. The original DOCEDI processed only a FBIS data base and, from it, produced an AFP-searchable file. Creation of an AFP search file requires special processing that adds AFP-recognizable codes to the documents in the file. For instance, beginning- and end-of-document codes and an end-of-search file code are inserted into the data base. Punctuation marks are separated from adjacent words by blank spaces.

This task involved modifying DOCEDI to accept messages in a GENSER format and processing them as just described to create an AFP searchable data base. The task was accomplished using the GENSER data collected as described in Section 3.3. Using this new GENSER feature of DOCEDI, it takes approximately fifty minutes to create a GENSER data base consisting of 20,000 blocks (over 10 million characters).

3.4.3 Appending Documents to Search Files. This task required amending DOCEDI to accept a single file and append it to an existing search file, which was accomplished.

3.5 Synonym Dictionary Capability in Queries

A Synonym Dictionary capability has been added to the AFP user interface software to facilitate the creation of more complex queries containing lists of synonym terms to be ORed together.

The Synonym Dictionary is a collection of synonym files, each containing a list of synonym terms. The files are created using the RSX-11D system editor (EDI). One or more of these lists may be included in a given query by referencing the Synonym Dictionary File name(s) within the query. The need for entering each synonym term is thereby eliminated.

The process for generating queries remains unchanged. Queries are created using the Query Language Translator in the on-line or indirect mode of operation. However, the QLT process is modified to recognize a Synonym Dictionary File name appearing in a user-entered query line or an indirect query record. A Synonym Dictionary File is signified by a file name preceded with an at-sign (@). When the software encounters the character "@", the Synonym Dictionary File is accessed and the records found in the Dictionary File are included in the query being built as ORed terms and phrases. When the end of the Dictionary File is reached, the file is closed, and QLT resumes processing the remainder of the current line or record.

The following is an illustration of the use of several synonym files in a single query. The example is designed to locate documents in a foreign broadcast data base which contain the Soviet reaction to the establishment of friendly relations (normalization) between The People's Republic of China and the United States.

The user enters:

@USSR 20 @USA 20 @PRC AND @NORMAL

where USSR, USA, PRC and NORMAL are Synonym Dictionary Files.

If these files contain the following:

USSR	USA	PRC	NORMAL
U . S . S . R .	U . S . A .	P . R . C .	VISIT
SOVIET UNION	UNITED STATES	CHINA	NORMALIZATION
RUSSIA	AMERICA	PEKING	TRADE
MOSCOW	WASHINGTON	BEIJING	ENVOY
TASS	NIXON	CHOU EN LAI	MISSION
IZVESTIA	U . S .	MAO TSE TUNG	DIPLOMATIC RELATIONS
BREZHNEV			

then the query interpreted by the Query Language Translator would read,

'U . S . S . R .' OR 'SOVIET UNION' OR 'RUSSIA' OR 'MOSCOW' OR
'TASS' OR 'IZVESTIA' OR 'BREZHNEV' AND WITHIN 20 WORDS
'U . S . A .' OR 'UNITED STATES' OR 'AMERICA' OR 'WASHINGTON' OR
'NIXON' OR 'U . S .' AND WITHIN 20 WORDS 'P . R . C .' OR
'CHINA' OR 'PEKING' OR 'BEIJING' OR 'CHOU EN LAI' OR 'MAO TSE TUNG' AND
'VISIT' OR 'NORMALIZATION' OR 'TRADE' OR 'ENVOY' OR 'MISSION' OR
'DIPLOMATIC RELATIONS'

The Synonym Dictionary File is created using the RSX-11 line editor. The only restrictions on the Synonym Dictionary File are:

(a) The file type must be ".SYN".

(b) Each term or contiguous-word phrase is entered on a single line

followed by a carriage term.

(c) No Boolean terms are included in the list.

(d) No punctuation may appear in the list.

3.6 Synonym Dictionary Listing Program

The purpose of the Synonym Dictionary Listing program (SYNDIC) is to provide a means of listing the contents of all of the synonym dictionary files on the line printer. These files contain query term synonyms which are indirectly introduced into the query string as described in Section 3.5.

SYNDIC lists each synonym file name along with the synonyms appearing in that file. It also lists all synonyms in the entire dictionary in alphabetical order with the name(s) of the file(s) in which they appear. It takes only a few minutes to list twenty or so synonym files.

3.7 Multi-User Capability on UNIVAC 1652

The purpose of this task is to enable four users to simultaneously access the AFP and run concurrent queries on UNIVAC 1652 display terminals. The 1652 allows the use of function keys for initiating query creation, document search and for retrieving documents. The dual screen capability of the 1652 also allows more document text to be displayed on the screen than on a single screen CRT terminal, since one 1652 screen is devoted to text, while the other is used to display the retrieval header containing retrieval statistics and document status.

The multi-user capability was installed on the AFP system in the NMIC PAS facility in Washington, D. C., where four UNIVAC 1652 terminals are available.

3.8 Distinguishing Among Responses to Multiple Queries

This feature is called Subqueuing and is an enhancement that was added to more fully exploit the multiple query search and retrieval capability of the AFP. It allows a user to distinguish among documents found as responses to different queries.

The capability for creating multiple queries and for retrieving documents based on a multi-query search existed prior to this contract. However, there was no way to selectively review just those documents responding to a specific query, since all of the hit documents, for a particular user, were placed in the same review queue. A user could relate a document to its corresponding query(s) only by retrieving the document at the terminal and viewing the query identification in the text header. Therefore, in order to review only those documents responding to a specific query, the user was forced to review all of the hit documents in his queue, or search against one query at a time.

Subqueuing allows a single user to distinguish among documents as responses to four (or fewer) unique queries. More than four queries are allowed; however, documents responding to the fifth and subsequent queries will be placed in the last subqueue.

New user commands have been incorporated into the retrieval software to per-

mit switching between subqueues. The user may skip to the next subqueue or go back to the previous one. The document header information, containing retrieval and document statistics, has been expanded to include the subqueue number.

3.9 Allowing for Larger Number of Query Responses (Overflow)

An Overflow Capability has been added to the AFP retrieval software to permit more documents to be retrieved.

Prior to this modification the number of documents that could be retrieved by a single user had been limited to one hundred fifty due to the memory requirements for storing document retrieval queues within the retrieval program memory. The number of documents that may be retrieved with the Overflow Capability is, for all practical purposes, unlimited. However, an upper limit may be pre-selected when the AFP software is assembled.

The Document Retrieval Monitor (AMON) was modified to operate with either the Overflow Capability, which creates document retrieval queues on disk, or with internal queues, as before. The mode of queuing is an assembly-time parameter.

With the exception of the added retrieval capability, the operation of the Overflow Capability is transparent to the user.

3.10 Printing a Document on the Line Printer

The capability for printing documents on a remote line printer has been added to the document retrieval software. This feature is called Document

Print and allows a user to direct selected documents to a line printer for hard copy, while reviewing documents retrieved by the AFP.

This is a multi-user function, operating asynchronously with the retrieval software. Users initiate document printing by a command from the terminal, and may direct as many documents to the printer as necessary. The printer will maintain a print queue for multiple print requests.

The only restriction on the Document Print function is that the document to be printed must appear on the terminal at the time the print command is given.

3.11 Creation of an Editable Document

The capability for creating an editable document from a search file document has been added to the AFP retrieval software.

This feature is called the Document Keep function and is invoked by a user command during document retrieval. After invoking Document Keep the user may continue reviewing documents.

Document Keep enables a user to create a file which may be opened for editing after the document retrieval session. With this feature documents may be corrected, updated, and appended. Reports containing documents or parts of documents may also be created.

3.12 Concordance (Synonym Candidate) Listing Program

The purpose of the Concordance Listing Program (SYNCAN) is to provide a line printer listing of all substantive words in a search file.

One of the primary uses of such a listing is to furnish a user with a list of words which might be considered candidates for synonyms to be placed in a synonym file. In addition, the listing provides the numbers of the documents in which a word is found and a count of the frequency of its occurrence in the search file. Approximately one hour is required to list a 2000-block (one million character) data base on a high speed line printer.

3.13 Deletion of Documents from Search File

The Document Delete feature is a file management utility that is used to remove unwanted documents from an archival disk file. Documents may be deleted for numerous reasons: if the document is invalid, untimely, or requires updating, for example.

The user interface to the document deletion function has been added to the Document Retrieval Monitor (AMON). The Document Delete function is invoked during document retrieval from the system manager's terminal (usually the system terminal). The document to be deleted must be displayed on the terminal when the deletion command is input. After the delete command has been given, no other retrieval functions may be performed until the deletion is completed.

When the Document Delete command is input, the Retrieval Monitor sends the current document address data to the Disk Edit Task (DSKEDI), then returns to wait for completion before returning control to the terminal. DSKEDI accesses the search file, overwrites with zeroes the disk blocks occupied by the document, then returns a completion code to the Retrieval Monitor.

Although not a contractual requirement, a File Compression function has been developed which operates in conjunction with File Deletion. File Compression is a utility which is used to compress out the empty disk blocks created by the File Deletion function. The compression of a file results in contiguous text blocks and additional empty blocks at the end of the file. The empty blocks may be used for new documents. The combined functions of the Delete and Compress utilities provide a search file maintenance capability whereby documents can be deleted and the empty space can be reused.

File Compression is initiated via the retrieval monitor; however, it should not be initiated while other users are attempting AFP search and retrieval.

The compression of a large file is a lengthy process (e.g., ten minutes for approximately five million characters); therefore, compression should be initiated only after a large number of file blocks have been deleted, or when the space at the end of the file for appending new documents is severely limited.

3.14 Address List for Performing Subset Queries

The Subset Address List is a disk file containing retrieval information for documents selected as responses to a query(s). The subset capability operates in conjunction with the Overflow feature described in subsection 3.9.

A separate Subset Address List file is created for each user; however, if the Subqueuing (subsection 3.8) feature is used, then a separate List is created for each unique query, for up to four queries per user. If a single

user searches against more than four queries, the fourth Subset Address List file will contain document pointers for the fourth and the additional queries.

A Subset Address List is created during an AFP search after a document hit has been sent to the retrieval software. This feature is transparent to the user during normal search and retrieval sessions.

The Lists are found in the AFP system user area (UIC), and are given names that identify the pertinent user and query. Each List has a name of the form SUBQxy, where x is the number of the originating terminal unit, and y is the subqueue number (0, 1, 2, or 3).⁹

The Subset Address List files belonging to a particular user are deleted when the next search is initiated by that user. Therefore, if the Lists are to be used later, they should be renamed prior to executing the next search.

3.15 Don't Care

This function was intended to allow a user to generalize a query by substituting "don't care" characters for words or characters in a query. The "don't care" characters would elicit a more comprehensive response to the query than their more specific counterparts.

Two different approaches were taken in an attempt to implement the "don't care" function on the AFP. Both were unsuccessful, however, because of an inherent hardware characteristic of the Associative Crosspoint Processor,

9. In the case where the Subqueuing assembly-time option is not used, the subqueue number will be zero in the Subset Address List file name.

the AXP. Because of this characteristic in the AXP, a problem of ambiguity arises when a "don't care" term is included in a multiple query search. After a "hit" on a "don't care" term, the Query Resolution software module cannot distinguish from the information returned by the AXP which terminal the hit is for.

There is still hope that an approach can eventually be discovered which will circumvent this ambiguity problem, but as of this writing the "don't care" function cannot be satisfactorily implemented, and this task is incomplete.

3.16 Highlighting

This capability permits a user to highlight (by increased intensity of the display) those words and/or contiguous-word phrases in a retrieved document which appeared as terms in his query. The user has the option of selecting for highlighting either all the terms in his query which are satisfied by the document or only those terms which actually caused the document to be selected as a "hit".

As of this writing, work on this task is continuing but incomplete. One reason for the incompleteness of this task is that considerable difficulties were encountered in trying to parallel in the software of the highlighting module the hardware functions performed by the AXP in establishing hits. These difficulties were finally resolved, however. Consequently, there is no reason to assume that the task cannot eventually be completed. Time on the referenced contract, however, ran out before the highlighting function could be successfully implemented.

4. AREAS FOR POSSIBLE FUTURE ENHANCEMENTS

4.1 General

Though the capabilities added to the AFP as a result of tasks performed under the referenced contract significantly enhance its usefulness in intelligence applications, several other improvements could be made to the AFP system to even further increase its versatility as an intelligence tool. Some of the areas in which further enhancements could be made are described below.

4.2 Receipt of Real-Time Messages

Modifications could be made to the AFP software to enable the AFP to receive real-time messages from GENSER or other sources and to build therefrom a search data base. The data base would be updated "on-the-fly" as each new message was received. In this way the AFP could be used with current and timely messages.

4.3 Automated Dissemination

This task would result in the development of an automated message dissemination system with a machine-aided manual distribution capability. The capability of this system to disseminate messages in a live traffic environment would be demonstrated. Live traffic could be simulated through the use of a representative data base stored on magnetic tape.

Incoming messages would be processed sequentially, formatted and stored on an AFP-searchable data file, then passed against user dissemination profiles

in the AFP. Messages matching one or more of the profiles would be queued for review at the appropriate analyst/work station. Messages not matching any dissemination profile would be placed in a supervisor review queue for subsequent disposition. Such messages would be manually designated by the supervisor for distribution to the appropriate queue, or would be printed or deleted from the queue. Dissemination profiles would be created in the same manner that queries are now created, and the profiles could be readily updated on-line.

4.4 Incorporation of PDP-11/04 into AFP System

This would provide for parallel processing of large, complex file systems. Search functions currently handled by the host PDP-11/45 would be transferred to the PDP-11/04. This would reduce the processing load on the PDP-11/45 and enable multiple AFP/PDP-11/04 modules to query, independently and in parallel, numerous large, dissimilar data bases. The task would entail the conversion of PDP-11/45 programs to PDP-11/04 format, and the development of various handlers and utility programs to accommodate the high-speed DMC channel which would be introduced into the system between the PDP-11/45 and the PDP-11/04.

4.5 Improved Editing and Report Generation

A CRT-screen-oriented editor would be developed to permit the creation and revision of messages and reports.

The editor would allow documents from the search file to be edited in a screen page format for subsequent report generation or for updating the

search file. New search documents and reports could also be created and edited in screen format.

The CRT-oriented editor would differ from the current PDP-11 RSX-11 editor, EDI, (a teletype, line-oriented editor), in that it would employ the features available in an intelligent terminal, such as DEC's VT100 or VT132, for scrolling forward and backward, positioning the edit cursor, and erasing lines and pages. Function keys would also be provided, allowing expanded capabilities.

The basic editor would allow RSX-11 files to be created and modified under editor control and would include the following features:

- Editor cursor control functions, e.g., up, down, right, left, home
- Text word, phrase, line and page deletion and insertion
- Text word or phrase finder
- Scroll forward and backward
- Move a text word or phrase
- Patch feature to replace each occurrence of a word or phrase with another
- Save and restore function
- File concatenation
- Report generation

Report generation would provide the following capabilities for creating formatted documents and professional reports: paragraphing, indentation, page numbering, and table, figure, title, and table-of-contents creation using special document-formatting macros at edit time.

4.6 Conversion to DEC IAS Operating System

This task would convert the current AFP software to run under the DEC IAS Operating System. The current software runs under Version 6.2 of the RSX-11D Operating System. By converting to IAS, the AFP software would be operating under a fully supported DEC product.

4.7 Conversion to DEC UNIX Operating System

The AFP software is currently written in Macro-11 Assembly Language to run under version 6.2 of the RSX-11D Operating System. With the growing acceptance of UNIX as a general-purpose, interactive, time-sharing operating system, it may be desirable to support the AFP in a UNIX environment.

There are numerous ways of approaching this problem. One method would be to implement the AFP search software and hardware on a smaller background processor, such as a PDP-11/04, having its own search disk, running under IAS or RSX-11M. The retrieval software could be rewritten to run under UNIX in the host processor, with interprocessor communications via a DMC.

If, however, additional hardware is not desirable, the AFP software could be rewritten to run under the UNIX Operating System.

The AFP is, in part at least, a real time system, requiring asynchronous I/O

and intertask communications, which are available under RSX-11D, RSX-11M and IAS. UNIX, however, does not directly support these requirements. The file structure and file I/O also differ under UNIX. RSX-11D, RSX-11M, and IAS are block-space and record-oriented, whereas UNIX file I/O is byte-oriented.

UNIX does not support the Macro-11 Assembly Language; therefore, all of the AFP software must be rewritten under another language. The primary language of UNIX is the "C" Language, which is a general purpose programming language. "C" possesses sufficient flexibility that it has displaced assembly language programming under UNIX. Therefore, "C" is a reasonable candidate for the AFP software under UNIX.

The successful implementation of the AFP software under the UNIX Operating System using "C" Language would require a phased approach.

Phase I would consist of analyzing the UNIX Operating System and "C" Language capabilities to determine where they directly support AFP functions and where redesign and special auxiliary software would be required. The result of Phase I would be a functional AFP system design that would be supported under UNIX and a software design of that system for the "C" Language.

Phase II would implement the AFP software in "C" Language to run under UNIX on a machine such as the PDP11/70.

Phase III would develop a large data base and a test which would demonstrate the query building, data base search, retrieval, and multi-user features of the AFP running under UNIX.

4.8 User Interface Enhancement

This task would simplify and streamline the existing procedures for entering queries, initiating searches, and executing all AFP functions. At present, an AFP user must possess some knowledge of the DEC RSX-11D Operating System procedures before he can operate the AFP. The results of this task would relieve him of the burden of acquainting himself with these RSX intricacies. Using easy-to-understand dialogue, the monitor developed under this task would lead the user step-by-step through an easy-to-follow, natural-language procedure for operation of all functions of the AFP.

A monitor would be developed with links to all AFP function programs. The monitor would display (on request) a menu to the user informing him of all the AFP functions available and the corresponding two-letter mnemonic to enter on the keyboard to initiate a particular function.

After entering the appropriate mnemonic, the desired AFP function program would execute, perhaps entering into its own dialogue with the user.

At the conclusion of the selected function program, the monitor would prompt the user to request another function (or the menu) or to exit the AFP system.

This task would require the design and implementation of an AFP monitor and the modification of all existing AFP function programs to communicate with the monitor.

4.9 Improved Response Time

This task would redesign the AFP software to speed up and optimize the query translation, AXP-memory-table building, and the query resolution processes. This redesign would improve the response time for user queries.

The current query translation and resolution processes require the use of the DEC-supplied Macro Assembler and the Task Builder. These two software packages perform the translation and compilation of the user queries into a tabular form manageable for query resolution. Since these two packages are general-purpose, there is a high system overhead with respect to elapsed time and memory requirements associated with their use. This results in unnecessarily long waits for a response by the user. To reduce the user wait time, the translation and resolution processes would be coded as internal AFP software functions. The overhead problems associated with the DEC software packages would thus be removed, increasing throughput for the user.

4.10 Search of a Mixed Format Data Base

This task would provide the AXP with a capability to search a mixed-format (fixed-field and free-form narrative) type record. With a mixed-format data record, user queries could search over the fixed field, the narrative, or both, simultaneously.

Several agencies have expressed an interest in a mixed-format type of record. These records consist of fixed-field parameters containing answers or responses to fixed questions or entries. The fixed field is then followed by a narrative entered by the analyst to elaborate on the fixed

parameters.

4.11 Use of AFP with ARPA Network

This task would launch a study effort to determine the capabilities of an AFP accessible through the ARPA network. One of the features of the AFP which could benefit an ARPA environment is its ability to support many users accessing a common data base. This feature could significantly reduce the load on a mini-computer system by providing for the processing of multiple queries from many users during the same search cycle. This would not result in any increase in the CPU execution time.

REFERENCES

- (1) Datafusion Corporation, File Management for AFP Systems, Concepts and User Instructions, Part Number UM131009, December 1978.
- (2) Operating Systems, Inc., User Operations Guide for the Associative File Processor (AFP), AA-GUIDE-000-001-1, OSI: D77-001, 20 September 1979.
- (3) Operating Systems, Inc., User's Manual for the Associative File Processor (AFP), OSI: AA-USERS-000-001-1, 20 September 1979.
- (4) Operating Systems, Inc., Associative File Processor Programmer's Manual, AA-PROG-000-000-1, OSI: D77-003, 13 April 1978.
- (5) Operating Systems, Inc., Document Delete Program Specification, 20 September 1979.
- (6) Operating Systems, Inc., Search File Compression Program Specification, 20 September 1979.
- (7) Operating Systems, Inc., Document Editor Programmer's Manual, PM131007 VO1R00, 31 July 1978.
- (8) Operating Systems, Inc., Concordance (Synonym Candidate) List Program (SYNCAN), 20 September 1979.
- (9) Operating Systems, Inc., Synonym Dictionary List Program (SYNDIC), 20 September 1979.



MISSION of *Rome Air Development Center*

RADC plans and executes research, development, test and selected acquisition programs in support of Command, Control Communications and Intelligence (C³I) activities. Technical and engineering support within areas of technical competence is provided to ESD Program Offices (POs) and other ESD elements. The principal technical mission areas are communications, electromagnetic guidance and control, surveillance of ground and aerospace objects, intelligence data collection and handling, information system technology, ionospheric propagation, solid state sciences, microwave physics and electronic reliability, maintainability and compatibility.